**10.1 TESTOVÁNÍ STATISTICKÝCH HYPOTÉZ (Přehled testů; Parametrické testy)**

Testování hypotéz je nejpoužívanější metoda na vyvozování závěrů o základním stat. souboru (populaci) na základě vlastností výběrového souboru (vzorku – „sample“).

Statistické hypotézy dělíme na  
- **parametrické** – závěry se týkají parametrů zákl. souboru (průměr, rozptyl, …)

- **neparametrické** – závěry se týkají jiných vlastností (tvaru rozdělení, závislosti proměnných, …)

Při testu statistické hypotézy se rozlišuje **nulová hypotéza H0** o níž má ***test*** rozhodnout zda se zamítne, či nikoliv a **alternativní hypotéza H1**., kterou přijmeme v případě zamítnutí nulové hypotézy.

Při rozhodování o přijetí, či zamítnutí hypotézy se můžeme dopustit jedné ze dvou chyb:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  | **V Ý S L E D E K T E S T U** | |
|  |  |  | **nezamítáme H0** | **zamítáme H0** |
|  |  |  | **(přijímáme H1)** |
| **S K U T E Č N O S T** | **platí H0** | | správně | chyba 1. druhu |
| pravděp. rozhodnutí = 1-α | pravděp. rozhodnutí = α |
| (spolehlivost) | (hladina významnosti) |
| **neplatí H0** | **(přijímáme H1)** | chyba 2. druhu  pravděp. rozhodnutí = β | správně  pravděp. rozhodnutí = 1-β (síla testu) |

Protože pravděpodobnosti α a β spolu souvisí, nelze obě minimalizovat. Zpravidla se tedy volí α. V praxi se požaduje její hodnota 0,05 (někdy, zejména ve zdravotnictví 0,01).

**V této kapitole se budeme zabývat testy pro výběry z normálního rozdělení.**

**!Poznámka:** Výběrový průměr:

ale výběrový rozptyl

***Odvození rozhodovacího pravidla***

Myšlenku odvodíme pro nejjednodušší případ: základní soubor má normální rozdělení s neznámou střední hodnotou *μ* (kterou chceme odhadnout), ale známým rozptylem *σ*2.

Testujeme hypotézu H0: *μ* = *μ*0 proti alternativě H1: *μ* ≠ *μ*0, kde *μ*0 je nějaké konkrétní číslo.

Provedeme tedy náhodný výběr *x*1, …, *xn* o rozsahu *n*. Vypočítáme výběrový průměr (jako bodový odhad *μ*). Pokud bude „hodně vzdálen“ od *μ*, tak H0 zamítneme. Hledáme tedy takové číslo *k*, že | – *μ* | > *k*, přičemž pravděpodobnost zamítnutí hypotézy H0, ačkoliv platí, je rovna předem zvolenému (malému) číslu α. Tedy:

P(*zamítáme* H0 | *platí* H0) = α

P(| – *μ*0| > *k* **|** *μ* = *μ*0) = α

P (A)

Z kapitoly „Náhodný výběr“ víme, že pokud má základní soubor rozdělení N(*µ*, *σ*2), tak výběrový průměr má rozdělení N(*µ*, ) a tedy náhodná veličina má při platnosti H0 normované normální rozdělení N(0, 1).

Podle definice kvantilu pro veličinu Z s rozdělením N(0, 1) platí: P, tedy i:

P (B)

Porovnáním vztahů (A), (B) vidíme, že: , z čeho po úpravě dostaneme:

Rozhodovací pravidlo testu je tedy:

H0 zamítáme (přijímáme H1), když , tj. když

Výrazu se říká testovací (též pivotová) statistika a budeme jej značit R. Konkrétní realizaci této statistiky na základě náhodného výběru budeme označovat .

Obor hodnot, pro které se H0 zamítá, se nazývá kritický obor; ten budeme značit W.

!! Tedy leží-li v kritickém oboru, tak zamítáme H0.

const: Solve[Integrate[1-c/x^2,{x,1,Infinity}]==1,c]

hustota pravděpodobnosti f(x): D[1 - c/x^2, x]

rozsah 1 ... 10, =2/(A1^3)

střední hodnota E(X): Integrate[x\*(2/x^3),{x,1,Infinity}]

***p - hodnota***

Jak bylo poznamenáno hned na začátku, hladinu významnosti α volíme vždy předem. Zvolíme-li např. α = 0,05 a na této hladině významnosti zamítneme H0, můžeme se ptát, jestli bychom H0 zamítli i na vyšší hladině, třeba α = 0,1.

Hraniční hodnota, tedy nejmenší hodnota, pro kterou bychom ještě H0 zamítli, se nazývá dosažená významnost, nebo častěji p-hodnota. V předešlém příkladu je to tedy hodnota *p*, pro kterou = . Aplikujeme-li opět definici kvantilu, dostaneme: P(X > ) =

1 – P(X ≤ ) =

1 – P(X ≤ ) =

1 – = a tedy: *p* = 2·

Je-li tedy *p* > α, tak H0 nezamítáme, je-li *p* ≤ α, tak H0 zamítáme na hladině významnosti α.

Často se ale používá tzv. čistý test významnosti, při kterém se rozhodnutí o (ne)zamítnutí H0 odvíjí pouze od hodnoty *p* takto:

*p* < 0,01 – zamítáme H0 (přijímáme H1)

0,01 < *p* < 0,05 – nerozhodná oblast (doporučuje se opakovat test s větším rozsahem výběru)

0,05 < *p* – nezamítáme H0

-------------------------------------------------------------

Ještě předtím, než přikročíme k jednotlivým testům, poznamenejme, že rozlišujeme tyto možnosti:

*(q je kvantil příslušného rozdělení)*

*FITko s jakou pravděpodobností %:*

*1 - CDF[PoissonDistribution[28], 30]*

1) hypotéza H0: např. μ1 = μ2 proti oboustranné alternativě μ1 ≠ μ2

Kritický obor W = *p* = 2·min{P(R ≤ ), P(R ≥ )}

2a) hypotéza levostranná H0: např. μ1 ≤ μ2 proti pravostranné alternativě μ1 > μ2

Kritický obor W = *p* = P(R ≤ )

2b) hypotéza pravostranná H0: např. μ1 ≥ μ2 proti levostranné alternativě μ1 < μ2

Kritický obor W =  *p* = P(R ≥ )

-------------------------------------------------------------

**JEDNOVÝBĚROVÉ TESTY**

součástky mm:

=NORM.INV(0.9,0,2)

**TEST HYPOTÉZY O ROZPTYLU (PŘI JEDNOM VÝBĚRU) – test o rozptylu**

(*μ*, *σ* – parametry zákl. souboru; , *s* – realizace čís. hodnot z výběrového souboru o rozsahu *n*)

Z jednoho výběru testujeme, jaký je rozptyl zákl. souboru (jehož *μ* neznáme) vzhledem k dané konstantě ():

(H0: = ; H1: ≠ ; nebo H1: < ; nebo H1: > )

Realizace pivotové statistiky: *r*0 = (kvantily budeme vybírat z (n – 1) )

Kritický obor pro: oboustranný test: W =

levostranný test: W =

pravostranný test: W =

**TEST HYPOTÉZY O STŘEDNÍ HODNOTĚ (PŘI JEDNOM VÝBĚRU)**

(*μ*, *σ* – parametry zákl. souboru; , *s* – realizace čís. hodnot z výběrového souboru o rozsahu *n*)

Z jednoho výběru testujeme, jaká je střední hodnota zákl. souboru vzhledem k dané konstantě (*μ*0):

(H0: *μ* = *μ*0; H1: *μ* ≠ *μ*0; nebo H1: *μ* < *μ*0; nebo H1: *μ* > *μ*0)

I.A) známe-li rozptyl σ2 zákl. souboru tzv. **jednovýběrový *z*-test**

Realizace pivotové statistiky: *r*0 = (kvantily budeme vybírat z N(0, 1))

Kritický obor pro: oboustranný test: W =

levostranný test: W =

pravostranný test: W =

I.B) neznáme-li rozptyl σ2 zákl. souboru tzv. **jednovýběrový *t*-test**

Realizace pivotové statistiky: *r*0 = (kvantily budeme vybírat z *t*(*n* – 1))

Kritický obor pro: oboustranný test: W =

levostranný test: W =

pravostranný test: W =

**DVOUVÝBĚROVÉ TESTY**

drát:

a) sum(12^k \* e^(-12)/k!, k=0 to 9)

b) exp(-0.6 \* 4)

**TEST HYPOTÉZY O ROVNOSTI ROZPTYLŮ (PŘI VÝBĚRU ZE DVOU SOUBORŮ)**

**(dvouvýběrový F-test)**

Máme dva základní soubory s neznámými parametry *μ*1, *μ*2; , . Pomocí dvou nezávislých náhodných výběrů: X11, X12, …, X1*n*1 o rozsahu *n*1 a X21, X22, …, X2*n*2 o rozsahu *n*2 chceme testovat hypotézu o vztahu mezi neznámými rozptyly.

H0:

Tedy: **A:** H0: ; H1: ≠

Realizace pivotové statistiky: *r*0 = (kvantily budeme vybírat z *F*(*nč* – 1; *nj* – 1))

Kritický obor pro oboustranný test: W =

**B:** H0: ; H1: <

Realizace pivotové statistiky: *r*0 = (kvantily budeme vybírat z *F*(*nj* – 1; *nč* – 1) !jiné pořadí č-j !)

(Tedy jiná pivot. st. než u oboustranného! – podíl bez ohledu na to, který rozptyl je větší)

Kritický obor pro levostranný test: W =

**C:** H0: ; H1: >

Realizace pivotové statistiky: *r*0 = (jako v B) (kvantily budeme vybírat z *F*(*nč* – 1; *nj* – 1) !jako v A!)

Kritický obor pro pravostranný test: W =

**TEST HYPOTÉZY O ROVNOSTI STŘEDNÍCH HODNOT (PŘI VÝBĚRU ZE DVOU SOUBORŮ)**

Máme dva základní soubory s  parametry *μ*1, *μ*2 (neznámými) a , (dále rozlišíme). Pomocí dvou nezávislých náhodných výběrů: X11, X12, …, X1*n*1 o rozsahu *n*1 a X21, X22, …, X2*n*2 o rozsahu *n*2 chceme testovat hypotézu *μ*1 – *μ*2 = c (zpravidla c = 0; tedy *μ*1 = *μ*2).

pokud bylo, s jakou pravděpodobností %:

1 - CDF[NormalDistribution[75, Sqrt[1500\*0.05\*0.95]], 90.5]

V podstatě tedy: H0: *μ*1 = *μ*2; H1: *μ*1 ≠ *μ*2; nebo H1: *μ*1 < *μ*2; nebo H1: *μ*1 > *μ*2)

**I.** známe-li rozptyly , tzv. **dvouvýběrový *z*-test**

Realizace pivotové statistiky: *r*0 = (kvantily budeme vybírat z N(0, 1))

Kritický obor pro: oboustranný test: W =

levostranný test: W =

pravostranný test: W =

**II.A)** neznáme-li rozptyly , , ale víme-li, že jsou stejné tzv. **dvouvýběrový *t*-test** při rovnosti rozptylů

Realizace pivotové statistiky: *r*0 = (kvantily budeme vybírat z *t*(*n*1 + *n*2 – 2) )

kde

Kritický obor pro: oboustranný test: W =

levostranný test: W =

pravostranný test: W =

**II.B)** neznáme-li rozptyly , . Navíc taky nejsou stejné tzv. **dvouvýběrový *t*-test** při nerovnosti rozptylů

Realizace pivotové statistiky: *r*0 = (kvantily budeme vybírat z *t*(*n*1 – 1) a *t*(*n*2 – 1) )

kde

Kritický obor pro oboustranný test: zamítáme H0

**Cohenův koeficient věcného účinku**

7. pták: 15 \* (0.7)^4 \* (0.3)^2 \* 0.7

z 10 ptáku více než 5: Sum[Binomial[10, k] \* (0.7)^k \* (0.3)^(10-k), {k, 6, 10}]

Při aplikaci dvouvýběrových t-testů si musíme uvědomit, že při velkých rozsazích náhodných výběrů může i malý rozdíl ve výběrových průměrech způsobit zamítnutí nulové hypotézy, ačkoliv z věcného hlediska tak malý rozdíl nemá význam. Naopak při malých rozsazích nemusí i značně velký rozdíl ve výběrových průměrech způsobit zamítnutí nulové hypotézy.

Ukazatelem praktické důležitosti rozdílu mezi výběrovými průměry je Cohenův koeficient věcného účinku:

a posouzení velikosti účinku je: ***d*** ***účinek*** *.*

do 0,2 zanedbatelný

0,2 – 0,5 malý

0,5 – 0,8 střední

nad 0,8 velký

**PÁROVÉ TESTY**

**TEST HYPOTÉZY O ROVNOSTI STŘEDNÍCH HODNOT (PŘI VÝBĚRU Z DVOUROZMĚRNÉHO**

**(párový t-test) ROZDĚLENÍ)**

Obecně jde tedy o situaci, kdy základní soubor představují dvě „spárované“ náhodné veličiny (obě měřené na stejných objektech) s neznámými středními hodnotami *μ*1, *μ*2 a rozptyly a .

Chceme testovat hypotézy o vztahu středních hodnot *μ*1, *μ*2 na základě dvourozměrného náhodného výběru (dvojic hodnot (X11, X21), (X12, X22), …(X1*n*, X2*n*)).

Myšlenka párového t-testu spočívá v tom, že nepočítáme zvlášť střední hodnoty pro vzorky X1 a X2, ale sledujeme, jestli je střední hodnota **rozdílu** středních hodnot vzorků X1 X2, rovna nule.

Formálně jde tedy o jednovýběrový t-test aplikovaný na náhodnou veličinu Y = (X1 – X2).

H0: *μ*Y = 0 *μ*(X1-X2) = 0 *μ*X1 – *μ*X2 = 0 *μ*X1 = *μ*X2

Tedy: (H0: *μ*1 = *μ*2; H1: *μ* ≠ *μ*0; nebo H1: *μ* < *μ*0; nebo H1: *μ* > *μ*0)

Realizace pivotové statistiky: *r*0 = (kvantily budeme vybírat z *t*(*n* – 1))

kde: ; ; ale

Kritický obor pro: oboustranný test: W =

levostranný test: W =

pravostranný test: W =

**TEST NEZÁVISLOSTI – PEARSONŮV KORELAČNÍ KOEFICIENT**

vagon

CDF[NormalDistribution[20, 5], 16]

1 - CDF[NormalDistribution[200, Sqrt[10]\*5], 235]

Opět máme „spárované“ náhodné veličiny (dvourozměrný zákl. soubor): (X1, Y1), (X2, Y2), …(X*n*, Y*n*)), ale tentokrát chceme rozhodnout, jsou-li (ne)závislé. (Přesněji nekorelované, ale u souboru s dvourozměrným normálním rozdělením je to totéž.)

Pomocí výběrové kovariance vypočítáme výběrový Pearsonův

korelační koeficient: *r*(X, Y) = = =

(H0: *ρ*(X, Y) = 0; H1: *ρ*(X, Y) ≠ 0; nebo H1: *ρ*(X, Y) < 0; nebo H1: *ρ*(X, Y) > 0)

Realizace pivotové statistiky: *r*0 = (kvantily budeme vybírat z *t*(*n* – 2))

Kritický obor pro: oboustranný test: W =

levostranný test: W =

pravostranný test: W =

Hypotézu o nezávislosti rovněž zamítáme ve prospěch oboustranné alternativy, když *r*(X, Y) neleží v intervalu: